AD-A056 762 TEXAS UNIV AT AUSTIN CENTER FOR CYBERNETIC STUDIES F/G 12/1
A SPECIAL PURPOSE LINEAR PROGRAMMING ALGORITHM FOR OBTAINING LE--ETC(U)
SEP 77 R D ARMSTRONG, E FROME N00014-75-C-0569
UNCLASSIFIED CCS-318 NL

1 OF 1
AD A056762

END
DATE FILMED
9-78
DDC

LEVEL II

12

B.S.

CENTER FOR CYBERNETIC STUDIES

The University of Texas
Austin, Texas 78712

LEVEL II

(11) (9) (14)

Research Report CCS-318

(6)

A SPECIAL PURPOSE LINEAR PROGRAMMING ALGORITHM FOR OBTAINING LEAST ABSOLUTE VALUE ESTIMATORS IN A LINEAR MODEL WITH DUMMY VARIABLES.

by

(10) Ronald D. Armstrong*
Edward Frome**

(11) September 1977 (12) 24p.

*The University of Texas at Austin, Austin, TX 78712

**Oak Ridge Associated Universities, Oak Ridge, TN 37830

(15) N00014-75-C-0569,
NSF-MCS707-00100

This research was partly supported by NSF Grant MCS707-00100, and ONR Contract N00014-75-C-0569 with the Center for Cybernetic Studies, The University of Texas. Reproduction in whole or in part is permitted for any purpose of the United States Government.

CENTER FOR CYBERNETIC STUDIES

A. Charnes, Director
Business-Economics Building, 203E
The University of Texas
Austin, Texas 78712
(512) 471-1821

1472

406197

78 07 24 005 LB

A SPECIAL PURPOSE LINEAR PROGRAMMING ALGORITHM FOR OBTAINING LEAST ABSOLUTE VALUE ESTIMATORS IN A LINEAR MODEL WITH DUMMY VARIABLES

Ronald D. Armstrong

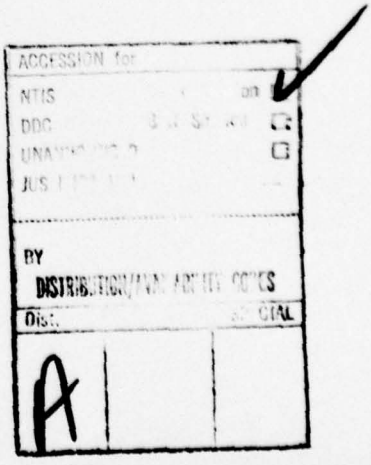
The University of Texas
Austin, Texas

Edward Frome

Oak Ridge Associated Universities
Oak Ridge, Tennessee

ABSTRACT

Dummy (0,1) variables are frequently used in statistical modeling to represent the effect of certain extraneous factors. This paper presents a special purpose linear programming algorithm for obtaining least-absolute-value estimators in a linear model with dummy variables. The algorithm employs a compact basis inverse procedure and incorporates the advanced basis exchange techniques available in specialized algorithms for the general linear least-absolute-value problem. Computational results with a computer code version of the algorithm are given.



1. INTRODUCTION

The standard linear regression model is usually written as follows:

$$y_i = x_{i1}\beta_1 + x_{i2}\beta_2 + \ldots + x_{im}\beta_m + \varepsilon_i, \quad i = 1, \ldots, n, \quad (1)$$

where $x_i = (x_{i1}, x_{i2}, \ldots, x_{im})$ are known values of the independent variables, y_i is an observed value of a dependent random variable, ε_i is an error term and the β_i's are the unknown parameters.

It is often desirable to include parameters in the model that represent the "effect" of different levels of one or more "factors". This can be done using dummy (0,1) variables to represent the levels of the factor, and is usually referred to as the analysis of covariance - e.g., see Searle (1971, chap. 4). Throughout this paper we assume that each observation is affected by at most one of the K levels of a single factor. The regression model then becomes

$$y_i = x_{i1}\beta_1 + x_{i2}\beta_2 + \ldots + x_{im}\beta_m + d_{i1}\alpha_1 + d_{i2}\alpha_2 + \ldots + d_{ik}\alpha_k + \varepsilon_i \quad (2)$$

where

$$d_{ik} = \begin{cases} 1 \text{ if } k\text{th level of factor is present,} \\ 0 \text{ otherwise, } k=1, \ldots, K, \end{cases} \quad (3)$$

and the α_k's are unknown parameters representing the effect of each level of the factor on the response.

The classical procedure for estimating the unknown parameters is to solve the following least squares problem.

$$\text{Minimize} \sum_{i=1}^{n} \left(y_i - \sum_{j=1}^{m} x_{ij}\beta_j - \sum_{k=1}^{K} d_{ik}\alpha_k\right)^2. \quad (4)$$

Computational techniques for solving (4) can be found in Searle (1971, chap. 4). Strong theoretical justification (Graybill (1961)) can be made for the least squares estimates when certain assumptions on the distributions of the random variables are made. However, in the presence of fat-tailed distributions or outliers, least absolute value (LAV) estimates may be recommended. Empirical studies (see Barrodale (1968), Glahe and Hunt (1970), Kiountouzis (1973)) comparing least squares and LAV estimators have demonstrated

the worth of the LAV criterion. The LAV problem may be stated:

$$\text{Minimize} \quad \sum_{i=1}^{n} \left| y_i - \sum_{j=1}^{m} x_{ij}\beta_j - \sum_{k=1}^{K} d_{ik}\alpha_k \right|. \qquad (5)$$

It is generally recognized that the LAV problem may be solved efficiently with a special purpose linear programming algorithm (see Barrodale and Roberts (1973)). The main purpose of this paper is to develop a further refinement of the linear programming approach to handle dummy variables. The advantages of the refinement are a reduction in computer processing time, and a reduction in computer storage.

Section 2 reviews the LAV algorithm of Barrodale and Roberts (1973) and develops the theory necessary to take advantage of the structure arising in the LAV problem when dummy variables are present. Section 3 discusses the computer code implementation and presents some computational results. Section 4 presents sample problems where LAV estimation is used.

2. THE LINEAR PROGRAMMING ALGORITHM

2.1 Review of the Barrodale and Roberts' Algorithm

We will begin this section with a brief review of a revised simplex implementation of the Barrodale and Roberts' algorithm. A more detailed description of this procedure - generalized to handle linear constraints - is given by Armstrong and Hultz (1977).

Charnes, Cooper and Ferguson (1955) demonstrate how the LAV problem associated with (1) can be written as the following linear programming (LP) problem.

$$\text{Minimize } z = \sum_{i=1}^{n} (P_i + N_i)$$

subject to

$$X\beta + IP - IN = Y$$

$$P \geq 0, \; N \geq 0$$

where X is the n by m matrix (x_{ij}), Y is the vector of y_i's, P and N are, respectively, the vectors of positive and negative deviations.

A basis for this LP problem is formed by m independent rows of

X. We denote this submatrix by X_B and its inverse by X_B^{-1}. We define IB to be the ordered set of row indices forming X_B (the row index of X corresponding to the first row of X_B is the first element of IB and so on) and NB is the index set of nonbasic rows.

We make the transformation $\lambda = X_B \beta$ and the problem becomes:

$$\text{Minimize } z = \sum_{i=1}^{n} (P_i + N_i)$$

subject to

$$\lambda_q + P_{i(q)} - N_{i(q)} = y_{i(q)}, \quad q = 1, 2, \ldots, m$$

$$X_i X_B^{-1} \lambda + P_i - N_i = y_i, \quad i \in NB$$

where i(q) is the q-th element of IB.

A current basic solution consists of

$$\bar{\lambda}_q = y_{i(q)}, \quad q = 1, 2, \ldots, m$$

$$\bar{\beta} = X_B^{-1} \bar{\lambda}$$

$$P_{i(q)} = N_{i(q)} = 0, \quad q = 1, 2, \ldots, m$$

$$\bar{P}_i = \begin{cases} y_i - X_i\bar{\beta};\ i \in NB,\ y_i - X_i\bar{\beta} > 0 \\ 0;\ i \in NB,\ y_i - X_i\bar{\beta} < 0 \end{cases}$$

$$\bar{N}_i = \begin{cases} X_i\bar{\beta} - y_i;\ i \in NB,\ X_i\bar{\beta} - y_i > 0 \\ 0;\ i \in NB,\ X_i\bar{\beta} - y_i \leq 0 \end{cases}$$

To assist in obtaining the LP reduced costs or rate of change achieved by removing a row from the basis, we define

$$\sigma_i = \text{sgn}\,(y_i - X_i\bar{\beta}),\ i \in NB$$

$$\sigma_i = +1 \text{ or } -1, \text{ when } y_i - X_i\bar{\beta} = 0,\ i \in NB.$$

When the deviation for a nonbasic row is zero, an initial assignment of +1 or -1 to σ_i is arbitrary and thereafter the algorithm will determine the value.

The LP dual variables (see Wagner (1959) for a statement of the dual problem) associated with the basic constraints are given by

$$\pi = (\pi_1, \pi_2, \ldots, \pi_m) = (\sum_{i \in NB} \sigma_i X_i) X_B^{-1}.$$

The dual variables associated with the nonbasic constraints are +1 or -1. Thus, the optimality condition for the specialized LP algorithm of Barrodale and Roberts is

$$-1 \leq \pi_i \leq +1, \; i = 1, 2, \ldots, m.$$

Let us assume that the optimality condition is not satisfied. Then, there exists a component of π (say, π_r) with $|\pi_r| > 1$. The row $X_{i(r)}$ has now been labeled as leaving the basis. The value of $\sigma = \text{sgn}(\pi_r)$ indicates whether λ_r is to be increased or decreased. Because the objective is to minimize the sum of the residuals, λ_r should be increased if π_r is negative and λ_r should be decreased if π_r is positive. The objective value may not strictly decrease at each iteration when degeneracy (i.e., $\bar{P}_i = \bar{N}_i = 0$, $i \in NB$) is present. Degeneracy never seems to cause cycling in practice and can be resolved with the perturbation technique of Charnes (1952). We will ignore the implications of degeneracy for the remainder of this paper.

Once the algorithm has specified that $X_{i(r)}$ is to leave the basis, it must determine the row of X to enter the basis in the r-th position. This is accomplished through a partial sort of the ratios:

$$(y_i - X_i\bar{\beta})/(\rho X_i X_{B(r)}^{-1}), \; i \in NB, \; \rho\sigma_i X_i X_{B(r)}^{-1} > 0$$

where $X_{B(r)}^{-1}$ is the r-th column of X_B^{-1}. Let L(u) denote the indices of $i \in NB$ forming the u smallest ratios. The algorithm then determines the value of u (say $\bar{u}$) that satisfies

$$|\pi_r| - \sum_{i \in L(\bar{u}-1)} \rho\sigma_i X_i X_{B(r)}^{-1} > 1$$

$$|\pi_r| - \sum_{i \in L(\bar{u})} \rho\sigma_i X_i X_{B(r)}^{-1} \leq 1.$$

Let t be the unique index which is in $L(\bar{u})$ and not in $L(\bar{u}-1)$. The row of X to enter the basis in the r-th position is X_t. The algorithm

updates X_B^{-1}, sets $\sigma_i = -\sigma_i$ for $i \in L(\bar{u}-1)$ and proceeds to the next iteration. The process terminates when $-1 \leq \pi_i \leq 1$, $i = 1, 2, \ldots, m$.

2.2 Extension of the LAV Algorithm to Handle Dummy Variables

Problem (5) can clearly be handled within the framework of the algorithm outlined in the previous subsection - we need only expand the observation matrix to include the dummy variables. However, we will show that substantial savings can be achieved by recognizing the problem's special structure.

The linear programming equivalent of problem (5) is:

$$\text{Minimize} \sum_{i=1}^{n} (P_i + N_i) \tag{6}$$

subject to

$$x_{i1}\beta_1 + x_{i2}\beta_2 + \ldots + x_{im}\beta_m + \sum_{k=1}^{K} d_{ik}\alpha_k + P_i - N_i = y_i;$$

$$P_i \geq 0,\ N_i \geq 0;\ i = 1, 2, \ldots, n.$$

Rewriting the constraints in matrix notation,

$$X\beta + D\alpha + IP - IN = Y.$$

Further simplifying the notation, we define

$$V = (X\ D) \text{ and } \gamma = \begin{pmatrix} \beta \\ \alpha \end{pmatrix}.$$

The constraints become

$$V\gamma + IP - IN = Y.$$

The algorithm of Barrodale and Roberts can now be applied directly with V and γ taking the place of X and β, respectively. Let V_B denote a basis for the problem formed by m + K independent rows of V. From the independence of the rows of V_B it follows that at least one V_i with $d_{ik} = 1$ (i.e., the m + k-th element of V_i equals 1) must be present in V_B for k = 1, 2, ...K. Hence, V_B may be partitioned as

$$V_B = \begin{pmatrix} X_F & D_F \\ X_G & I \end{pmatrix}$$

where X_F is an m by m submatrix of X, X_G is a K by m submatrix of X, I is a K by K identity matrix and D_F is an m by K matrix with at most one nonzero entry in each row (the nonzero entry being unity).

The inverse of V_B is given by

$$V_B^{-1} = \begin{bmatrix} (X_F - D_F X_G)^{-1} & -(X_F - D_F X_G)^{-1} D_F \\ -X_G(X_F - D_F X_G)^{-1} & I + X_G(X_F - D_F X_G)^{-1} D_F \end{bmatrix}.$$

We will demonstrate how the steps of the Barrodale and Roberts' algorithm can be executed conveniently with existing data, $W = (X_F - D_F X_G)^{-1}$ and a knowledge of the rows forming X_F and X_G. The logic behind this refined algorithm will be similar in many respects to the generalized upper bounding techniques (see Dantzig and Van Slyke, (1967)) of linear programming. In fact, upon taking the LP dual of (6), generalized upper bound constraints can be recognized.

Proceeding in a manner analogous to that used in subsection 2.1, we define NB to be the index set of nonbasic rows and i(q) to be the q-th basic row. We also define Y_F and Y_G to be the vector with components y_i corresponding to X_F and X_G.

Associated with V_B, we have a current solution

$$\begin{pmatrix} \bar{\beta} \\ \bar{\alpha} \end{pmatrix} = V_B^{-1} \begin{pmatrix} Y_F \\ Y_G \end{pmatrix}.$$

Through the previous partitioning scheme,

$$\bar{\beta} = WY_F - WD_F Y_G$$

$$\bar{\beta} = W(Y_F - D_F Y_G)$$

$$\bar{\beta} = W(Y^*),$$

where the q-th component of Y* equals $y_{i(q)} - y_{i(m+k)}$ when $d_{i(q),k} = 1$, $k = 1, 2, \ldots, K$ and equals $y_{i(q)}$ when $d_{i(q),k} = 0$, $k = 1, 2, \ldots, K$. Now α_k is easily determined by

$$\bar{\alpha}_k = y_{i(m+k)} - X_{i(m+k)}\bar{\beta}.$$

The sign of the residual for the nonbasic rows may be obtained in a similar manner to that given in subsection 2.1. The only difference is an extra addition that is required to evaluate the residual.

The value of the vector of dual variables at any iteration can also be calculated easily from W. We have

$$\pi = (\pi_1, \pi_2, \ldots, \pi_m, \pi_{m+1}, \ldots, \pi_{m+K}) = \Big(\sum_{i\in NB} \sigma_i(X_i, D_i)\Big)V_B^{-1},$$

where D_i is the i-th row of D.

By utilizing the partitioning of V_B^{-1},

$$\pi = \Big(\sum_{i\in NB} \sigma_i(X_iW - D_iX_GW), \quad \sum_{i\in NB} \sigma_i(D_i + D_iW_GWD_F)\Big).$$

Let $d^* = \sum_{i\in NB} \sigma_iD_i$ and $Z = \sum_{i\in NB} \sigma_iX_i - d^*X_G$, then $\pi = (ZW, d^* - ZWD_F)$.

Finally, we define the index set

$$Q(k) = \{q \mid d_{i(q),k} = 1;\ k = 1, 2, \ldots, K;\ q \leq m\}$$

It can now be observed that the first m components of π are equal to ZW and the last K components are given by

$$\pi_{m+k} = d_k^* - \sum_{q\in Q(k)} \pi_q,\ k = 1, 2, \ldots, K.$$

The optimality condition is

$$-1 \leq \pi_i \leq +1,\ i = 1, 2, \ldots, m+K.$$

Assuming the optimality condition is not satisfied, we have $|\pi_r| > 1$ and $\rho = \text{sgn}\ (\pi_r)$. Three mutually exclusive and exhaustive cases can arise.

CASE 1. $r > m$, $d_{i(r),\bar{k}} = 1$ and $Q(\bar{k}) = \phi$. *This means that the row* to leave the basis corresponds to the only observation in the basis affected by factor $\bar{k}$. Since the new basis must be nonsingular, the row to enter the basis has the form (X_i, D_i), $d_{ik} = 1$. The ratio test is simplified to consider

$$|y_i - X_i\,\bar{\beta} - \bar{\alpha}_k|;\ d_{ik} = 1,\ i\in NB,\ \rho\sigma_i > 0.$$

Let L(u) denote the indices of i yielding the u smallest of these values. Then $\bar{u} = [\pi_r/2]$, where $[\pi_r/2]$ indicates the greatest integer less than $\pi_r/2$. Let t be unique index which is in $L(\bar{u})$ and not in

$L(\bar{u}\text{-}1)$. The row (X_t, D_t), enters the basis in the r-th position. Notice that W does not change and the algorithm may proceed directly to the next iteration after updating σ_i, $i \in L(\bar{u}\text{-}1)$.

CASE 2. $r \leq m$. The ratios are given by

$$(y_i - X_i\bar{\beta} - \bar{\alpha}_k)/(\rho(X_i - X_{i(m+k)})W_r);\ i \in NB,\ d_{ik} = 1,\ k = 1, 2, \ldots, K$$

$$\rho\sigma_i(X_i - X_{i(m+k)})W_r > 0$$

and

$$(y_i - X_i\bar{\beta})/(\rho X_i W_r),\ i \in NB;\ d_{ik} = 0,\ k = 1, 2, \ldots, K;\ \rho\sigma_i X_i W_r > 0$$

where W_r is the r-th column of W.

The remaining steps are completely analogous to those given previously and will not be repeated.

CASE 3. $r > m$, $d_{i(r),\bar{k}} = 1$ and $Q(\bar{k}) \neq \phi$

In this situation we reorder the rows of V by interchanging

$$(X_{i(r)}, D_{i(r)})$$

with an $(X_{i(r^*)}, D_{i(r^*)})$, $r^* \in Q(\bar{k})$ (in other words, $d_{i(r^*),\bar{k}} = 1$). The new W denoted by W* is given by

$$W^*_{qj} = W_{qj} \qquad q \neq r^*,\ j = 1, 2, \ldots, m$$

and

$$W^*_{r^*j} = \sum_{q \in Q(\bar{k})} W_{qj},\quad j = 1, 2, \ldots, m.$$

We are now in CASE 2 with $r = r^*$ and may proceed as previously indicated.

3. IMPLEMENTATION AND COMPUTATIONAL RESULTS

Both algorithms described in section 2 have been coded in FORTRAN by the authors. They are maintained as independent subroutines with all input and output as parameters of the CALL statements. We refer to the subroutine corresponding to the revised simplex implementation of Barrodale and Roberts as L1NORM. The subroutine corresponding to the algorithm outlined in the subsection 2.2 is called L1DUM. The remainder of this section compares these subroutines with regard to computer and solution times.

3.1 Computer Storage.

Subroutine L1NORM saved the V matrix explicitly. This could easily be reduced with extra coding and without knowledge of the compact inverse procedures discussed here. The same can not be said of the basis inverse. This was represented as an (m+K) by (m+K) matrix in L1NORM and as an m by m matrix in L1DUM.

L1DUM maintained a capacity of $K \leq 15$, $n \leq 300$ and $m \leq 15$. While L1NORM maintained a capacity of $m+K \leq 25$ and $n \leq 300$. Some additional coding was, of course, required for L1DUM. Overall L1NORM utilized approximately 1500 more words of internal storage than L1DUM.

We did not employ double precision accuracy on any variables. On some machines (we ran on a CDC 6600 with a sixty bit word) or when solving less stable problems, double precision might be recommended and then the difference in storage utilization would be more dramatic. It should be noted that although we have developed the algorithms using an explicit representation of the inverse, the algorithm can easily be adapted to other methods for solving linear systems. In particular, decomposition methods such as those discussed by Bartels and Golub (1969) and Cline (1976) may be recommended in certain instances.

3.2 Solution Time.

A battery of test problems were solved with two computer codes L1NORM and L1DUM. L1DUM and L1NORM began with the same initial basis and used the rule of the max $\{|\pi_i|\}$ to define the row to leave the basis at any iteration.

Table I presents a summary of the computational results. Problem sets were randomly generated with five problems in each set. All runs were on the CDC 6600 at The University of Texas at Austin. Times are CPU times in seconds. The calls to the system clock were made in the main program immediately before and immediately after the respective subroutines were called. The number of pivots indicates the number of times the basis inverse was updated.

TABLE I

PROBLEM SIZE				L1DUM		L1NORM	
#	K	m	n	Time	Pivots	Time	Pivots
1	4	4	60	.103	13.8	.119	15.0
2	8	4	60	.105	13.4	.186	17.6
3	12	4	60	.119	15.8	.263	20.2
4	4	8	60	.210	20.4	.216	20.4
5	8	8	60	.212	22.4	.304	24.2
6	12	8	60	.213	23.6	.429	26.6
7	4	4	120	.210	13.0	.235	15.2
8	8	4	120	.314	20.4	.447	25.4
9	12	4	120	.298	19.2	.598	28.0
10	4	8	120	.549	28.6	.506	29.8
11	8	8	120	.578	32.8	.709	35.4
12	12	8	120	.566	31.6	.898	36.0
13	4	4	180	.450	19.2	.491	22.6
14	8	4	180	.550	22.4	.766	29.4
15	12	4	180	.476	20.6	.948	33.2
16	4	8	180	.894	31.2	.826	32.0
17	8	8	180	1.052	37.8	1.139	39.2
18	12	8	180	1.128	42.0	1.634	49.8
19	4	4	240	.613	29.2	.609	20.2
20	8	4	240	.885	26.2	1.227	35.2
21	12	4	240	.890	30.4	1.399	38.6
22	4	8	240	1.326	35.2	1.230	37.2
23	8	8	240	1.492	40.2	1.608	43.0
24	12	8	240	1.758	49.0	2.291	54.4
25	4	4	300	.995	24.4	.998	25.6
26	8	4	300	1.083	26.4	1.408	33.6
27	8	4	300	1.125	27.6	1.952	42.6
28	4	8	300	1.997	43.2	1.768	43.4
29	8	8	300	2.104	45.4	2.256	49.4
30	12	8	300	2.423	53.6	3.191	62.4

This table presents computational results with thirty sets of test problems. Five problems were solved for each dimension and reported times and iterations are means of the results. All times are in CPU seconds.

L1DUM performed better in most instances. As expected, the relative efficiency of L1DUM increased as the number of parameters associated with dummy variables (i.e., the value of K) increased.

4. APPLICATIONS

In problems that require dummy variables, the effect of outliers becomes more difficult to assess as the number of independent variables and dummy variables increases. The motivation for using a robust procedure such as LAV estimation is the same for problems that require dummy variables as it is in the general regression situation. The main purpose of this paper is the refinement of the linear programming approaches to LAV estimation. In this section, we discuss two applications of LAV estimation that involve dummy variables, and a small numerical example is presented.

There are several ways that dummy variable techniques are used in statistical data analysis. One situation occurs when a designed experiment is carried out, and the experimenter is aware of one or more "concomitant variables" that may affect the response and cannot be controlled by the experimenter. This situation is usually referred to as the analysis of covariance and combines features of regression analysis and analysis of variance.

To illustrate the estimation procedure, we use the data in Table II. The response is the average daily gain of pigs, and the experimenter is interested in the effect of four feeds. There are two covariates - initial age and weight - which are believed to affect weight gain. The values of the two covariates and the dummy variables are given in Table II. The LAV estimates of the parameters are given in Table III, which also shows the least squares estimates. If the experimenter wants to evaluate the relative importance of one of the covariates, he can delete that column and recompute the LAV estimates. It is possible to develop various approaches - see e.g., McNeil and Tukey (1975) - to assessing the "quality of fit" associated with a given model that can be used in a manner that parallels

TABLE II SOURCE: Snedecor and Cochran (1967, p. 440)

Days x_{i1}	Pounds x_{i2}	Pounds/Day y_i	Feed 1 d_{i1}	Feed 2 d_{i2}	Feed 3 d_{i3}	Feed 4 d_{i4}
78	61	1.40	1	0	0	0
90	59	1.79	1	0	0	0
94	76	1.72	1	0	0	0
71	50	1.47	1	0	0	0
99	61	1.26	1	0	0	0
80	54	1.28	1	0	0	0
83	57	1.34	1	0	0	0
75	45	1.55	1	0	0	0
62	41	1.57	1	0	0	0
67	40	1.26	1	0	0	0
78	74	1.61	0	1	0	0
99	75	1.31	0	1	0	0
80	64	1.12	0	1	0	0
75	48	1.35	0	1	0	0
94	62	1.29	0	1	0	0
91	42	1.24	0	1	0	0
75	52	1.29	0	1	0	0
63	43	1.43	0	1	0	0
62	50	1.29	0	1	0	0
67	40	1.26	0	1	0	0
78	80	2.67	0	0	1	0
83	61	1.41	0	0	1	0
79	62	1.73	0	0	1	0
70	47	1.23	0	0	1	0
85	59	1.49	0	0	1	0
83	42	1.22	0	0	1	0
71	47	1.39	0	0	1	0
66	42	1.39	0	0	1	0
67	40	1.46	0	0	1	0
67	40	1.36	0	0	1	0
77	62	1.40	0	0	0	1
71	55	1.47	0	0	0	1
78	62	1.37	0	0	0	1
70	43	1.15	0	0	0	1
95	57	1.22	0	0	0	1
96	51	1.48	0	0	0	1
71	41	1.31	0	0	0	1
63	40	1.27	0	0	0	1
62	45	1.22	0	0	0	1
67	39	1.36	0	0	0	1

LAV Estimation For Covariance Model

significance testing in the LS analysis. There is, however, no sampling theory for L1NORM estimates. Consequently, these model-building techniques are most useful in exploratory data analysis.

TABLE III

	Parameter					
	β_1	β_2	α_1	α_2	α_3	α_4
LAV	-.006706	.009594	1.466	1.326	1.310	1.310
LS	-.003454	.007414	1.337	1.182	1.318	1.217

Estimates of the parameters for data in Table II.

A second situation in which dummy variables can be used is when adjusting for the effect of seasonal factors in an economic time series. Suppose that the trend-cycle portion of the time series is represented with an "empirical function" composed of polynomial pieces called cubic splines, and that y_i is the observed value of the time series in month i, for i=1, ..., n. Then the x_{ij}'s in equation (2) are defined as follows:

$$x_{ij} = i^j, \ (j = 1, 2, 3),$$

$$x_{ij} = \begin{cases} (i-t^*_{j-3})^3 \text{ if } i \geq t^*_{j-3}, \ (j = 4, \ldots, m+3), \\ 0 \end{cases}$$

$$x_{ij} = \begin{cases} 1 \text{ if } i - K\,[(i-1)/K] = j - (m+3), \ (j = m+4, \ldots, K+m+4) \\ 0 \text{ otherwise,} \end{cases}$$

where the knots $\{t^*_j;\ j = 1, \ldots, m\}$ are known constants which divide the time domain into m+1 intervals, and there are k = 12 levels of the seasonal factor. Frome and Armstrong (1977) considered LAV estimation for this model and have presented a numerical example with m = 9 and n = 132 where the response is the residential construction authorized in Texas for each month in the time period from 1966 to 1976.

5. CONCLUSIONS

With the development of more efficient computer codes (algorithms) to obtain the L_1 norm estimates, the L_1 norm has become an important tool in data analysis. This paper has demonstrated how the Barrodale and Roberts' algorithm can be specialized to solve a class of problems involving dummy variables which arise in several instances. The specialized algorithm has been shown to reduce solution times and computer storage requirements when several columns of the model are associated with dummy variables. Although we have not studied the numerical properties in detail, numerically stable procedures for solving linear systems can easily be combined with the framework of the algorithm.

The two FORTRAN subroutines used to obtain the results reported here are available from the authors for a handling charge.

BIBLIOGRAPHY

Armstrong, R. D. & Frome, E. L. (1976). A comparison of two algorithms for absolute deviation curve fitting. J. Amer. Statist. Assoc. 71, 328-30.

Armstrong, R. D. & Hultz, J. W. (1977). A restricted discrete approximation problem in the L_1 norm. SIAM J. Numer. Analysis 14, 555-65.

Barrodale, I. (1968). L_1 Approximation and the analysis of data. Appl. Statist 17, 51-7.

Barrodale, I. & Roberts, F. D. K. (1973). An improved algorithm for discrete L_1 linear approximation. SIAM J. Numer. Analysis 10, 839-48.

Barrodale, I. & Roberts, F. D. K. (1974). Solution of an overdetermined system of equations in the L_1 norm. Commun. Assoc. Comput. Mach. 17, 319-20.

Bartels, R. H. & Golub, G. H. (1969). The simplex method of linear programming using LU decomposition. Commun. Assoc. Comput. Mach. 12, 266-8.

Charnes, A. (1952). Optimality and degeneracy in linear programming. Econometrica 20, 160-70.

Charnes, A. & Cooper, W.W. (1961). Management Models and Industrial Applications of Linear Programming, Vol. I and II. New York: John Wiley and Sons, Inc.

Charnes, A., Cooper, W.W., & Ferguson, R. (1955). Optimal estimation of execution compensation by linear programming. Management Sci. 2, 138-51.

Cline, A.K. (1976). A descent method for the uniform solution to overdetermined systems of linear equations. SIAM J. Numer. Analysis 13, 293-309.

Dantzig, G.B. & Van Slyke, R.M. (1967). Generalized upper bounding techniques. J. Computer Systems Sci. 1, 213-26.

Frome, E.L. & Armstrong, R.D. (1977). A robust procedure for estimating the trend-cycle component of an economic time series. Proc. 10th Ann. Symp. Comput. Sci. Statist. Interface (in press).

Glahe, F.R. & Hunt, J.G. (1970). The small sample properties of simultaneous equation least absolute estimators vis-a-vis least squares estimators. Econometrica 38, 742-53.

Graybill, F.A. (1961). An Introduction to Linear Statistical Models. New York: McGraw-Hill.

Kiountouzis, E.A. (1973). Linear programming techniques in regression analysis. Appl. Statist. 22, 69-73.

McNeil, D.R. & Tukey, J.W. (1975). Higher order diagnosis of two-way tables, illustrated on two sets of demographic empirical distributions. Biometrics 31, 487-510.

Searle, S.R. (1971). Linear Models. New York: John Wiley and Sons, Inc.

Snedecor, G.W. & Cochran, W.G. (1967). Statistical Methods. Iowa State University Press.

Spyropoulos, K., Kiountouzis, E. & Young, A. (1973). Discrete approximations in the L_1 norm. The Computer J. 16, 180-86.

Wagner, H.M. (1959). Linear programming techniques for regression analysis. J. Amer. Statist. Assoc. 54, 206-12.

Unclassified

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author)	2a. REPORT SECURITY CLASSIFICATION
Center for Cybernetic Studies The University of Texas	Unclassified 2b. GROUP

3. REPORT TITLE

A Special Purpose Linear Programming Algorithm for Cbtaining Least Absolute Value Estimators in a Linear Model with Dummy Variables

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

5. AUTHOR(S) (First name, middle initial, last name)

Ronald D. Armstrong
Edward Frome

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
September 1977	18	20

8a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S)
NSF Grant MSC707-00100; N00014-75-C-0569 b. PROJECT NO.	Center for Cybernetic Studies Research Report CCS 318
c.	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)
d.	

10. DISTRIBUTION STATEMENT

This document has been approved for public release and sale; its distribution is unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
	Cffice of Naval Research (Code 434) Washington, D. C.

13. ABSTRACT

Dummy (0, 1) variables are frequently used in statistical modeling to represent the effect of certain extraneous factors. This paper presents a special purpose linear programming algorithm for obtaining least-absolute-value estimators in a linear model with dummy variables. The algorithm employs a compact basis inverse procedure and incorporates the advanced basis exchange techniques available in specialized algorithms for the general linear least-absolute-value problem. Computational results with a computer code version of the algorithm are given.

DD FORM 1 NOV 65 1473 (PAGE 1)

S/N 0101-807-6811

Unclassified

Security Classification

A-31408

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Least Absolute Values						
L_1 Norm						
Analysis of Covariance						
Regression						
Generalized Upper Bounding						
Linear Programming						